Tc MRKSVCPKQKFFFSAFPFFFFFCVFPLISRTGQEKLLFDQKYKIIKGEKKEKKKNQRANRREHQQKREIKRFKKS   76

Tc FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIEEGADLGMVF   151

Tc MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVVLDTPAGLVRGTAHLQSGTESEVSNASIINVPSFLYQ   226

Tc QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC   300

Tc VEIYGPPTNPEANYKNVVVIFGNRQADR SPCGT GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE   371

Tc   RIPGVKVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*   423

```
Tc                                    RTGQEKLLFDQKYKIIKGEKKEKKKNQPANRREHQQKREIMRFKKS            75

Tc    FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAPLFDPIEEGADLGMVF    150

Tc    MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVVLDTPAGLVRGTAHLQSGTESEVSNASIINVPSFLYQ    225

Tc    QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC    300

Tc    VEIYGPPTNPEANYKNVVIFGNRQADR | SPCGT | GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE    371

Tc    RIPGVKVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*        423
```

87

Tc   MRKSVCPKQKFFFSAFPFFFFFCVFPLIS

SUBSTITUTE SPECIFICATION

SEQUENCE ID NO.4

[illegible] 75

[illegible] 150

[illegible] 225

[illegible] 300

[illegible] SPGT [illegible] 375

[illegible] 423

```
Cs   --------------------------------------------------------------KNFSKG                            6

Cs   IHAIDSHTMGEPTRIVVGGIPQINGETMADKKKYLDNLDYVRTALMHEPRGHNDMFGSIITSSNKKEADFGIIF                        81

Cs   MDGGGYLNMCGHGSIGAATVAVETGMVEMVEPVTNIN--MEAPAGLIKAKVMVEN---EKVKEVSITNVPSFLYM                      151

Cs'  EDAKLEVPSLNKTITFDISFGGSFFAIIHAKELGVKVETSQVDVLKKLGIEIRDLINEKIKVQHPELEHIKTVDL                      226

Cs   VEIYDEPSNPEATYKNVVIFGQGQVDR  [ SPCGT ]  GTSAKLATLYKKGHLKIDEKEVYESITGTMFKGRV--LEET             297

Cs   KVGEFD-----------AIIPEITGGAYITGENHEVIDPEDPLKYGFTV--- 335
```

SEQ ID NO:6

```
Pa    --------------------------------------------------------------MQR       3

Pa    IRIIDSHTGGEPTRLVIGGFPDLGQGDMAERRRLLGERHDAWRAACILEPRGSDVLMGALLAPVDPEACAGVIF     73

Pa    FNNSGYLGMCGHGTIGLVASLAHLGRIGPGV------HRIETPVGEVEATLH-------EDGSVSVRNVPAYRVR    140

Pa    RQVSVEVPGI-GRVSGDIAWGGNWFFLVAGH--GQRLAGDNLDALTAYTVAVQQALDD----QDIRGEDGGAIIH    213

Pa    IELFAD--DPHADSRNFVLCPGKAYDR SPCGT GTSAKLACLAADGKLLPGQPWRQASVIGSQFEGRYEWLDGQ    279


Pa     PGGPIVPTIRGRAHVSAEATLLLADDDPFAWGIRR- ----------------    314
```

SUBSTITUTE SPECIFICATION

Splice leader
acceptor sites

Signal Peptide

Polypyrimidine rich region
⇩

```
CCTTTTTCTTTTTAAAAACAAAAAAAATTCCGGGGGGAATATGGAACAGGGTATATGCGTAAAGTGTCTGTCCCAAACAAAATTTTT      90
                                                       M  R  K  S  V  C  P  K  Q  K  F  F     12
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTTTTGTGTGTGTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTGACCAAAATAT     180
 F  S  A  F  P  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  D  Q  K  Y     42
AAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAGGGAAATTATGCGATTT     270
 K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  E  H  Q  Q  K  R  E  I  M  R  F     72
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT     360
 K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N     102
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT     450
 M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F     132
GGAGCCTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT     520
 G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H     162
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA     610
 N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T     192
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCGTCATT     700
 P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F     222
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT     790
 L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I     252
GTTCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT     880
 V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N     282
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCA     970
 R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A     312
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT     1060
 N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y     342
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA     1150
 A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R     372
ATACCCGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG     1240
 I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M     402
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT     1330
 G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  *                            423
GGGAACAGGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTATTTTTTATTTTTATTAGTGCATT     1420
ATTATTAAATTTTTTTTTTGTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGAAGCGATAGCGGCCGGACAATTTTTTGTTTTAT     1510
TTTCATTTTCATCTTCCTACCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGGTCTTGTCGGTGGAGGAGTCCTAAATCCCGCACCTCGG     1600
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT     1651
```

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(d)   Nucleotide sequence and peptide sequence TcPA45

```
                                              M  R  K  S  V  C  P  K  Q  K  F  F    180
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTGTGTGTTTCCTTGATCTCTCGAACAGGGCAGGAAAGCTTCTGTTTGACCAAAAT    42
  F  S  A  F  P  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  D  Q  K  Y    270
AAAATTATTAAGGGCGAGAAAAAGAAAAGAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAGGGAAATTATGCGATT    72
  K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  E  H  Q  Q  K  R  E  I  M  R  F    360
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT    450
  K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N    430
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT    132
  M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F    520
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT    162
  G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H    610
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA    192
  N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T    700
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATT    222
  P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F    790
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTTCGCCATT    252
  L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I    850
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT    282
  V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N    970
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCA    312
  R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A    1060
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT    342
  N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y    1150
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA    372
  A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R    1240
ATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTATCATG    402
  I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M    1330
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT    423
  G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  *                            1420
GGGAACACGTCCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTATTTTTATTTTTATTTTATTAGTGCATT    1510
ATTATTAAATTTTTTTTTTTGTTTTGGGGTTTCAACGGTACGGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGCTTTTAT    1600
TTTCATTTTCATCTTCCTACCCAACCCCCTTGGTTCCACCGGTCGTGGCGGGGTCTTGTGGGTGGAGGAGTCCTAAATCCCGCACCTGG    1651
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT
```

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(a)   Nucleotide sequence and peptide sequence TcPA45

SUBSTITUTE SPECIFICATION

```
                                      CGAACAGGGCAGGAAAAGCTTCTGTTTGACCAA AAATAT  270
 F  S  A  F  P  F  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  Q  K  Y    72
AAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAGGGAAATTATGCGATTT  360
 K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  E  H  Q  Q  K  R  E  I  M  R  F   102
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT  432
 K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N   132
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT  522
 M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F   162
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT  610
 G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H   192
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA  700
 N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T   222
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATT  790
 P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F   252
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT  880
 L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I   282
GTTCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT  970
 V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N   312
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCGGAGGCA 1060
 R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A  342
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT 1150
 N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y  372
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA 1240
 A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R  402
ATACCGGGGGTGAAGGTGCCGGTGACCAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG 1330
 I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M  423
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT 1420
 G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  -
GGGAACACGTGCAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTATTTTTATTTTTATTTTTATTAGTGCATT      1510
ATTATTAAATTTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTGCTTTTAT  1600
TTTCATTTTCATCTTCCTAGCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGGTCTTGTGGTGGAGGAGTCCTAAATCCGGCACCTCGG  1651
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT
```

┌─────────────────────────┐
│ Polyadenilation site   │
└─────────────────────────┘

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

┌──┐
│ (b)   Nucleotide sequence and peptide sequence TcPA45   │
└──┘

94

SUBSTITUTE SPECIFICATION

Signal peptide

M R K S V C P K Q K F F

Nucleotide sequence of signal sequence TcPA45

95

SEQUENCE ID NO:11

ATGCGATTT<

[illegible]

SEQ ID NO : 12

5'TTICCRAADATIACIACGTT 3'

SEQ ID NO : 13

5' ATHGCITTYGGIGGIAAYTTT 3'

SEQ ID NO : 14

5' TTICCRAADATIACIACGTT 3'

SEQ ID NO : 15

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'

SEQ ID NO : 16

5' CTGAGCTCGACCAGATCTATCTGC 3'

SUBSTITUTE SPECIFICATION

```
   1 cctttttctt tttaaaaaca aaaaaaattc cgggggggaat atggaacagg gtatatgcgt
  61 aaaagtgtct gtcccaaaca aaaatttttt ttttccgcct tcccattttt ttttttttt
 121 tgtgtgtttc ccttgatctc tcgaacaggg caggaaaagc ttctgtttga ccaaaaatat
 181 aaaattatta agggcgagaa aaaagaaaag aaaaaaaatc aacgagcaaa caggagagaa
 241 caccaacaaa aaacgggaaat tatgcgattt aagaaatcat tcacatgcat cgacatgcat
 301 acggaaggtg aagcagcacg gattgtgacg agtggtttgc cacacattcc aggttcgaat
 361 atggcggaga agaaagcata cctgcaggaa aacatggatt atttgaggcg tggcataatg
 421 ctggaaccac gtggtcatga tgatatgttt ggagccttt tatttgaccc tattgaagaa
 481 ggcgctgact tgggcatcgt attcatggat accggtggct atttaaatat gtgtggacat
 541 aactcaattg cagcggttac ggcggcagtt gaaacgggaa ttgtgagcgt gccggcgaag
 601 gcaacaaatg ttccggttgt cctggacaca cctgcgggt tggtgcgcgg tacggcacac
 661 cttcagagtg gtactgagag tgaggtgtca aatgcgagta ttatcaatgt acccccatt
 721 ttgtatcagc aggatgtggt ggttgtgttg ccaaagccct atggtgaagt acgggttgat
 781 attgcattg gaggcaattt ttcgccatt gtcccgcgg agcagttggg aattgatatc
 841 tccgttcaaa accctccag gctgcaggag gcaggagaac ttctgcgtac tgaaatcaat
 901 cgcagtgtga aggttcagca ccctcagctg ccccatatta acactgtgga ctgtgttgag
 961 atatacggtc cgccaacgaa cccggaggca aactacaaga acgttgtgat atttggcaat
1021 cgccaggcgg atcgctctcc atgtgggaca ggcaccagcg ccaagatggc aacactttat
1081 gccaaaggcc agcttcgcat cggagagact tttgtgtacg agagcatact cggctcactc
1141 ttccagggca gggtacttgg ggaggagcga ataccggggg tgaaggtgcc ggtgaccaaa
1201 gatgccgagg aagggatgct cgttgtaacg gcagaaatta ctggaaaggc ttttatcatg
1261 ggtttcaaca ccatgctgtt tgacccaacg gatccgttta agaacggatt cacattaaag
1321 cagtagatct ggtagagcac agaaactatt gggaacacg tgcgaacagg tgctgctacg
1381 tgaagggtat tgaatgaatc gtttttttt atttttatt tttatttta ttagtgcatt
1441 attattaaat tttttttttg tttgggggtt tcaacggtac cgcgttggga gcagggaagc
1501 gatagcggcc ggacaatttt ttgctttat tttcatttc atcttcctac ccaaccccct
1561 tggttccacc ggtcgcggcg gggtcttgtg ggtggaggag tcctaaatcc cgcacctcgg
1621 aggaataaac atatttcaat ttcatatctt ggaatcaaaa ggcat
```

SEQ ID NO : 18

WIIK

SEQ ID NO : 19

IVTGSLPDISG

SEQ ID NO : 20

ATNVPVVLDTPAGLVR

SEQ ID NO : 21

VDIAFGGNF

SEQ ID NO : 22

NVVIFGNR

SEQ ID NO : 23

MATLYAK

SUBSTITUTE SPECIFICATION

SEQ ID NO : 24

5' TCCGTATCCATGTCGATGC 3'

SEQ ID NO : 25

5' TATTATTGATACAGTTTCTG 3'

SEQ ID NO : 26

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'

SUBSTITUTE SPECIFICATION